

AT&T Speech Technologies and Applications

Jay Wilpon
Executive Director, Speech Services Research
AT&T Labs

October 30, 2009

Evolution of Communications Services

- Limited access
- Limited communication devices; personalize services
- People intense services
- User Experience was natural



1900
Voice



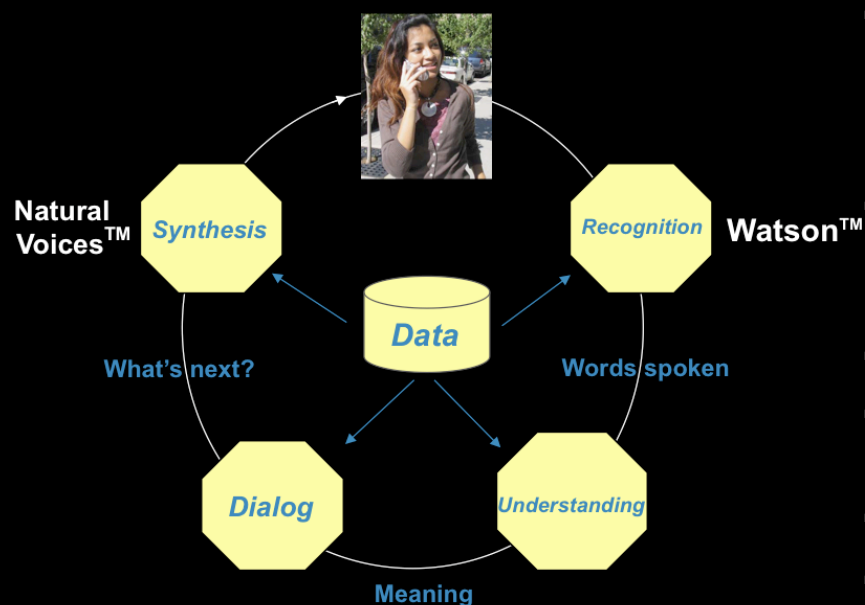
1970
Touch Tone / DialTone



"2009"
Multimodal / Data

AT&T Speech Technologies and Services

- AT&T Speech Technologies have been successfully **deployed for 25 years**
- **Hundreds of voice-enabled applications** deployed
- Deployed the industry's first nationwide voice-enabled service **saving AT&T > \$3B**
- Deployed the industry's first natural language understanding based service (AT&T VoiceTone®)



- AT&T is **recognized worldwide** for its pioneering research in speech recognition (**Watson**), text-to-speech synthesis (**Natural Voices**), natural language understanding, machine learning, speaker verification
- **Invented speaker independent ASR**, word spotting, barge-in, spoken natural language understanding, first 1M word speech recognizer
- **Co-created Voice XML** standard for IVR services
- Over 1M hours of research investment over three decades
- Nat'l Academy of Sciences members, IEEE Fellows
- 400+ US Patents Issued; 4000+ published papers
- AT&T's Research innovations are used in most commercial speech products and services

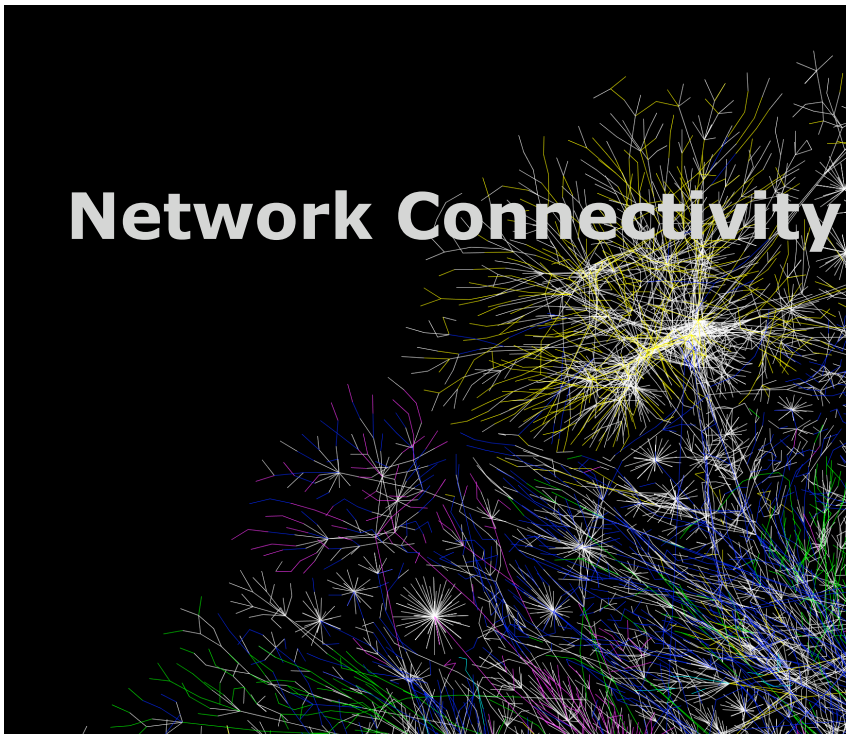


Key Opportunities for Speech Technologies

- **Increased service quality** – Easier access to information & services: education, government services, healthcare
- **Increased access to information** – Better accessibility to information & services for disabled and elderly populations
- **Cost savings** – drive more efficiency in healthcare and government services
- **Security** – increase security for people and property; secure access to information and services
- **Safety** – safe mobile environment; hands busy – eyes busy



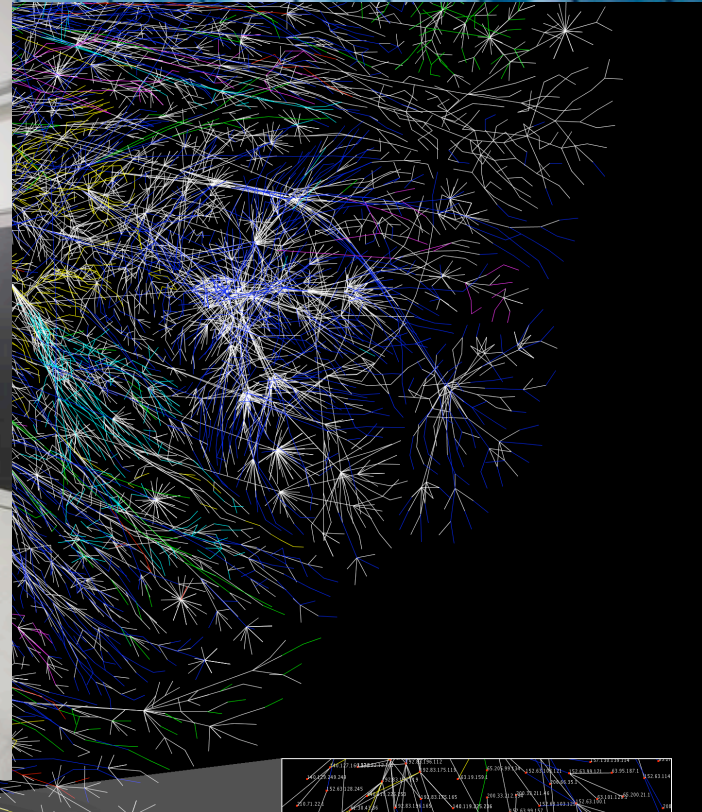
Network Connectivity



Cloud Computing



Bandwidth



"Driving" Innovation: Speech in the Cloud

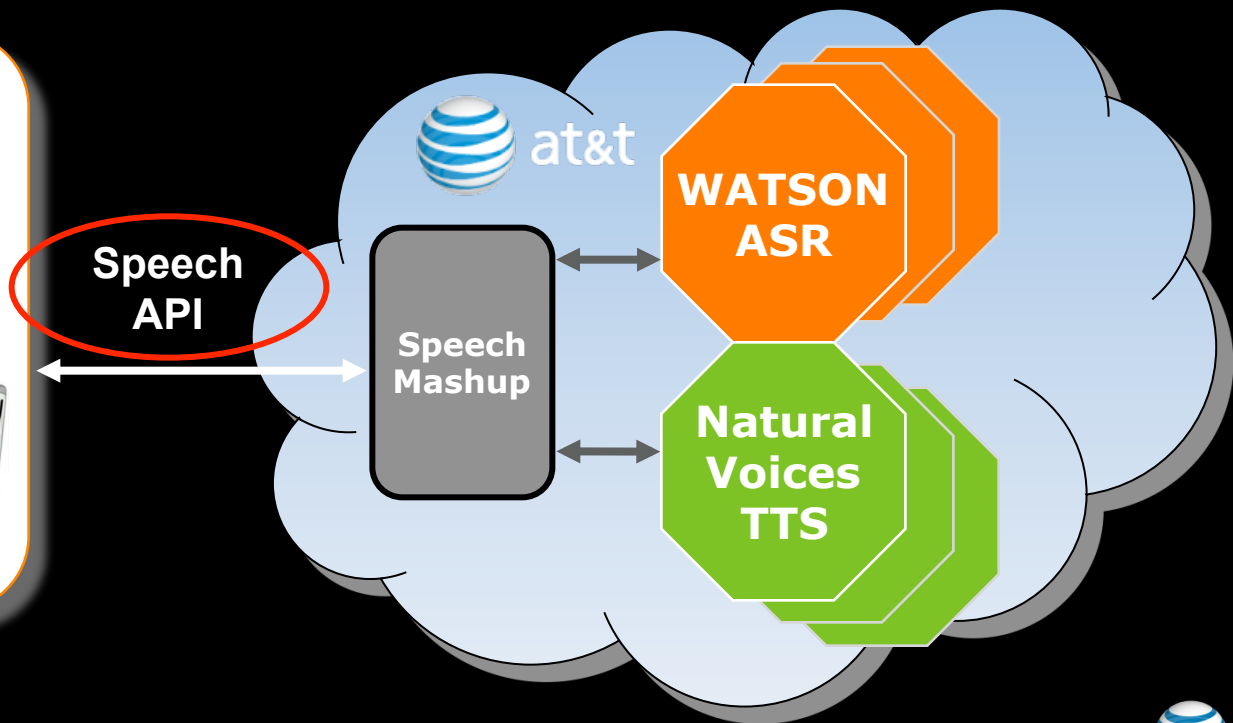
Grow development ecosystem; Enable easy creation of Multimodal Services

Enable Best in Breed: Provide speech technologies and APIs, thus reducing high barrier-of-entry for 3rd party developer ecosystem and enabling a new generation of service innovators

Benefit from Best in Breed Service Innovators: Many new services, many new vendors, increase customer base, new business models



Devices



“Opportunities”

Speak4it – Information Search

Say It! Get It!
Speak4it!
The easiest way to find things
in your local area.

Available on the iPhone
App Store

speak4it
MOBILE LOCAL SEARCH

pizza in san francisco

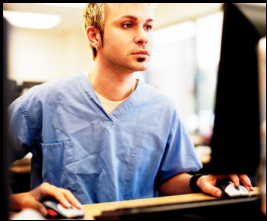
Go Getters Pizza	0.2 mi
609 Gough St, San Francisco	
Vincio Pizzeria	0.2 mi
201 Ivy St, San Francisco	
S. F. Pizza	0.2 mi
1270 Market St, San Francisco	
Pomodoris Pizzas	0.3 mi
1704 Market St, San Francisco	
Stacks	0.4 mi
11 Hayes St, San Francisco	
Papa's Chicago Pizza	0.4 mi
511 Hayes St, San Francisco	

Tap

Powered by at&t.

Healthcare and Speech Technologies (1/2)

Challenge



Senior physicians poorly equipped for electronic medical records



Poor patient compliance leads to avoidable, expensive interventions

Technology

Large-vocabulary **speech recognition** for dictation and search of medical information

Spoken dialog system calls patients to confirm meds taken; alerts doctors to irregular monitor readings

Benefit

Better adoption of electronic medical records by key senior doctors

Healthier patients; alerts doctors to problems earlier



Healthcare and Speech Technologies (2/2)

Challenge

Technology

Benefit



Poor access to care for patients with limited English abilities

Automatic translation for patient-doctor interaction

Better patient communication; lower interpretation costs



HIPAA: enable trusted but efficient access to records

Voice biometrics to authenticate patients and providers

Multi-factor verification for increased security; obviates need for passwords which can be lost or stolen

SmarTV – Language Translation



Mobile Devices – Anytime, Anywhere Access



Assistive Services and Speech Tech (1/2)

Challenge

Technology

Benefit



People who are blind need access to text & Web-based information

Recognition, language understanding and synthesis for web browsing, & transactions; ebook interactions; TV access

Access to education, work
Privacy



People with low vision can't use regular maps, smart phones; GPS devices, can't read product info

Spoken dialog for local search, navigation, messaging
Bar code readers + speech for shopping

Easier access to info and travel
Inexpensive or free

iWalk – Walking Direction for Low Vision People



Assistive Services and Speech Tech (2/2)

Challenge

Technology

Benefit



People who are deaf need help to communicate via phone, watch TV

Text-to-speech for dual party relay
Speech recognition for closed captioning

Access to communication



People with mobility impairments can't easily get to the office or school

Office-in-a-box for remote work
Speech recognition for command+control

Autonomy